



Evidence for the Identification of Carabayo, the Language of an Uncontacted People of the Colombian Amazon, as Belonging to the Tikuna-Yurí Linguistic Family

Frank Seifart^{1*}, Juan Alvaro Echeverri²

¹ Max Planck Institute for Evolutionary Anthropology, Leipzig, Germany, ² Universidad Nacional de Colombia, Amazonia Campus, Leticia, Colombia

Abstract

This paper provides evidence for the identification of the language of the uncontacted indigenous group called Carabayo, who live in voluntary isolation in the Colombian Amazon region. The only linguistic data available from this group is a set of about 50 words, most of them without reliable translations, that were collected in 1969 during a brief encounter with one Carabayo family. We compare this material with various languages (once) spoken in the region, showing that four attested Carabayo forms (a first person singular prefix and words for ‘warm’, ‘father’, and ‘boy’) display striking similarities with Yurí and at least 13 Carabayo forms display clear correspondences with contemporary Tikuna. Tikuna and Yurí are the only two known members of the Tikuna-Yurí linguistic family. Yurí was documented in the 19th century but has been thought to have become extinct since. We conclude that the Carabayo – directly or indirectly – descend from the Yurí people whose language and customs were described by explorers in the 19th century, before they took up voluntary isolation, escaping atrocities during the rubber boom in the early 20th century.

Citation: Seifart F, Echeverri JA (2014) Evidence for the Identification of Carabayo, the Language of an Uncontacted People of the Colombian Amazon, as Belonging to the Tikuna-Yurí Linguistic Family. PLoS ONE 9(4): e94814. doi:10.1371/journal.pone.0094814

Editor: Mark Aronoff, Stony Brook University, United States of America

Received: January 29, 2014; **Accepted:** March 20, 2014; **Published:** April 16, 2014

Copyright: © 2014 Seifart, Echeverri. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The authors have no support or funding to report.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: frank_seifart@eva.mpg.de

Introduction

There are still around 100 uncontacted indigenous groups around the world, and a few dozen of them in the Amazonian rainforest, according to the NGO Survival International (<http://www.survivalinternational.org/tribes>). Most of these groups are known to be closely related linguistically and culturally to groups already contacted. However, not much more than their mere existence is known about some of them. This is the case for the Carabayo people who live in the remote upper River Puré and River Bernardo region in the Colombian Amazon rainforest. The name Carabayo derives from the (mock) name “Bernardo Caraballo”, which was given to a Carabayo man by local people during a brief encounter in the Colombian town La Pedrera (Bernardo Caraballo was the name of a Colombian boxing champion). Subsequently the Carabayo people and their language have been referred to as Caraballo [1,2]. The 2013 *Ethnologue* language catalogue [3] introduced Carabayo as an English version of the language name, and assigned the ISO 639-3 code cby to it.

In the current study, we analyze the only linguistic data available from this group, around 50 words that were overheard and noted down during a brief encounter with one Carabayo family in 1969, showing that the Carabayo most likely speak a language closely related to Yurí (also spelled Jurí) as well as to Tikuna, which have previously been shown to be related to each other [2,4–7]. The Yurí language was documented in four wordlists in the 19th century but had been presumed to have

become extinct since. Tikuna is still spoken by about 40,000 speakers, mostly along the Amazon River in Peru, Colombia, and Brazil. If Carabayo is relatively closely related to both of these languages, as we suggest here, one possibility would be that it may be part of a former dialect continuum circumscribed by the Tikuna and Yurí languages.

Our study substantiates previous claims of a link between Carabayo and Yurí that were either based on limited data and non-rigorous methods [8] or did not substantiate this claim with linguistic data at all [9,10]. We also substantiate the existence of similarities between Carabayo and Tikuna that were noted by Goulard & Montes Rodriguez [2] based on incomplete Carabayo materials which they considered to be too poor to draw any further conclusions. Our identification is based on a comparison of all available Carabayo data (from three different sources) with, firstly, four Yurí wordlists collected in the early to mid-19th century, one of which has only recently become accessible [11]; Natterer’s Yurí wordlist was thought to have been destroyed in a fire in Vienna in 1848, until it was discovered in the late 1970s by Ferdinand Anders in the University Library of Basle. The handwritten manuscript has recently been transliterated by Hélène B. Brijnen at Leiden University. (Incidentally, the Carabayo wordlist [12] was also not accessible [13] until recently, because the Capuchin missionary publication *Amanecer Amazónico*, in which it was published, was not distributed widely. Additionally, the issue of *Amanecer Amazónico* that contains the Yurí wordlist is missing in both the Capuchin missionary archive in Leticia and in the

national library of Colombia, the *Biblioteca Luis Ángel Arango* in Bogotá. It was eventually located by the first author in the *Arxiu Provincial dels Caputxins de Catalunya* in Sarrià, Barcelona.) Secondly, we compare Carabayo with contemporary Tikuna data provided by a native speaker of Tikuna. Our identification of the relationship of Carabayo with Yuri and Tikuna also implies that Tikuna should no longer be considered a language isolate with no living relatives [3].

The nature and scarcity of the available Carabayo data implies that standard methods for identifying languages – e.g. by frequent sequences of sounds or letters [14] – or for establishing genealogical relations between languages – e.g. by regular sound changes in sets of cognate words [15] – cannot be applied straightforwardly. Our analysis of the available Carabayo data thus draws on a variety of methods to derive evidence for the likelihood of an identification of Carabayo. These include phoneme frequency counts, semantic extensions of words, taking into account the context in which Carabayo words occurred, morphological composition of words, and the relative borrowability of different sections of vocabulary.

The Carabayo material investigated here was collected in 1969 from people who live in the upper River Puré/upper River Bernardo area, between the Putumayo and Caquetá rivers in the Colombian Amazon region [16]. In early 1969, a local Colombian and a local Miraña Indian undertook an expedition to the Carabayo's territory. When they did not return, a military commission that was sent to rescue them made violent contact with the Carabayo people and took one family hostage. This family, consisting of an adult couple and three children, was then held in the boarding school of the Capuchin mission in the Colombian town La Pedrera for a few weeks before they were 'repatriated'. During this encounter, the Carabayo data analyzed here were collected. Since then, Carabayo people and traces of them have been sighted on various occasions, primarily by members of the cocaine mafia and guerilla fighters [16]. The most recent evidence of the Carabayo's persistence are aerial photographs of their roundhouses taken in 2010 [16].

We strongly disapprove of the circumstances under which Carabayo data were collected. We hope that our study of the Carabayo material that was published in reports of these dire events contributes to the protection of the Carabayo people, in line with, for example, Survival International's policy of making knowledge about uncontacted peoples public in order to raise awareness of the threats they are facing. In 2002, the Río Puré National Park was created to protect the Carabayo's territory. In addition, a legal decree passed in 2011 guarantees uncontacted peoples in Colombia such as the Carabayo the rights to their voluntary isolation, their traditional territories, and reparations if they face violence from outsiders.

Since the language of the Carabayo was unintelligible to any of the indigenous peoples present in La Pedrera in 1969, it is a mystery as to who the Carabayo are. Trupp [9] and Landaburu [10] have hypothesized that the Carabayo people might be descendants of the Yuri, without, however, discussing linguistic evidence [13]. They were apparently not aware of Vidal y Pinell's [8] attempt to analyze the Carabayo data published by Juan Berchmans de Felanix [17]. (Note that names of Capuchin monks are composed of a religious name followed by the place of their origin and that they are ordered alphabetically by the religious name, following the conventions established by the *Lexicon Capuccinum* [18]. The secular name of Juan Berchmans de Felanix was Antonio Font.) Vidal y Pinell concluded that the language of the Carabayo corresponds to Yuri, based on two arguments. Firstly, he suggests that three items from a Yuri wordlist by

Wallace [19] correspond to items collected from Carabayo in 1969 by Juan Berchmans de Felanix [17]. Secondly, Vidal y Pinell [8] compared the frequencies of the sound *tʃ*, which he considered the most "representative" phoneme in Wallace's Yuri data, in Carabayo, Yuri, and various other languages spoken in the region, for which Wallace [19] also provides wordlists. He found that it occurs in 23% of the words in Wallace's Yuri word list (18/77) and in 21% of the words from Juan Berchmans de Felanix's Carabayo list (7/33), but in maximally 8% of the words of Kubeo, Tucano, Kueretú, Tariana, and Baniwa. From these two pieces of evidence, Vidal y Pinell [8] concludes that the Carabayo that were sighted in 1969 are descendants of the Yuri, documented by Wallace around 1850. The current study support the hypotheses of a genealogical link between Carabayo and Yuri, based on a much more detailed discussion of potentially cognate forms. Crucially, this discussion is based not only on Yuri data collected by Wallace but also on Yuri data collected by Spix, Martius, and Natterer. In addition, we include correspondences with contemporary Tikuna in our discussion.

Materials and Methods

For the purpose of the current study, all attested linguistic material reported for the Carabayo family in 1969 has been gathered (Table 1). Most of it is from the list published by Juan Berchmans de Felanix [17]. A few additional items are interspersed in two descriptions of the encounter with the Carabayo in 1969: One by Juan Berchmans de Felanix himself [20] and another by Venanci d'Arenys de Mar [21]. These items include four clearly Spanish words that the Carabayo reportedly used but that were apparently learned during the brief encounter, i.e. *tabako* (item 1) and *karaβayo to comes* from item 27, which probably correspond to the Spanish-based name given to the Carabayo man by the people of La Pedrera and to Spanish *tu comes* 'you eat'. Excluding these four words, there are a total of 55 word tokens. Within these, one word occurs twice (*uro*, item 6), two three times (*na*, items 3, 7; *kariba*, items 3, 5, 28), and one four times (*ane*, after merging *n* and *nn*, see below, items 29, 30, 35), i.e. there are 48 word types in the data. Two elements in the list are of Nheengatu (Tupian) origin, which was the lingua franca used in the area in the 17th, 18th and 19th centuries. These are *kariba* 'white man' (items 3, 5, 28), and *tupana* 'God' (item 23).

Juan Berchmans de Felanix [17] notes that the translations he provides are very hypothetical, in fact mere guesses, given that he and the Carabayo had no common language. Juan Berchmans de Felanix was a native speaker of Catalan and fluent in Spanish. His publication with the Carabayo vocabulary was written in Spanish and meant for a Spanish-speaking readership. We thus assume that the phonetic values of the consonants and vowels in his representation of Carabayo correspond to those of Spanish graphemes and the Carabayo material was transliterated to IPA symbols accordingly. Additionally, *x* (item 21) was transliterated as *ʃ*, following the pronunciation rules of Catalan, Juan Berchmans de Felanix's native language, since [ʃ] has no graphic representation in Spanish. We transliterated *ê* as *ə* based on his remark that it stands for "ê neutra francesa [neutral French ê]" [17]. An acute accent appears in only four items (7, 9 = 18, 10, and 29 = 30). This suggests that whatever it may represent, it was not systematically marked. Therefore we disregard it for establishing hypothetical phoneme inventory.

Tables 2 and 3 are hypothetical phoneme charts of Carabayo with indications of phoneme frequencies in the extant data. Some aspects of this hypothetical phonology must remain uncertain because a few putative phonemes occur only once or twice (in

Table 1. All attested linguistic material for Carabayo,

	Carabayo	gloss, explanation, or context	source
1	tabako	'tobacco'	[20]
2	hako	at being frightened by dogs; 'bite' according to Juan Berchmans de Felanix [17]	[17,20]
3	'jja kariba, 'jja irobe	shouted at white people by an old women during the occupation of her house. Castro Caycedo [30] reports that Carabayo contacted on a path shouted <i>kariba</i> , <i>kariba ñé</i>	[20]
4	jje	'no'	[17,21]
5	kariba	'white man'	[17,21,30]
6	uro, uro	when meeting a white man in the bush, pointing at direction opposite to where he came from	[20]
7	jja-nauué	'give me, show me'	[17]
8	gudda	'wait'	[17]
9	agó	'bring'	[17]
10	amá	'come' (Spanish <i>sigá</i>)	[17]
11	ao	how the children call their father	[17]
12	aúa	calling a child	[17]
13	gu	'yes'	[17]
14	hono	'boy'	[17]
15	jja	'out', maybe based on item 03	[17]
16	pama	'there, look!'	[17]
17	pinə	'shrimp/prawn'	[17]
18	pinə-gó	'bring shrimp/prawn' (see items 09, 17)	[17]
19	tʃauameni	'good, well, like'	[17]
20	tʃaunoʃe	'warm me!' (the speaker ordered a child to warm his hands with fire and apply them to his body)	[17]
21	ʃama	'enough!'	[17]
22	alo	'come!'	[17]
23	tupana	'God'	[21]
24	jakoma	boy's name; according to Bergès [26] the autodenomination of the Carabayo is <i>yacumo</i> .	[21]
25	jakomanate	man's name	[21]
26	jego	'child(ren)', used by Carabayo woman addressing (two of) her children	[21]
27	oro kami karaβayo to comes	'give me meat, Carabayo wants to eat'	[21]
28	kariba dimene	(during forced walk through jungle), <i>dimene</i> means 'kill' according to Juan Berchmans de Felanix [17]	[17,20]
29	ané ui konə	/	[17]
30	ané uikarə	/	[17]
31	arə tʃe o neko	/	[17]
32	bajaneku	/	[17]
33	ekoneko	/	[17]
34	enənənəna pikhu	/	[17]
35	erə anne anne	/	[17]
36	etamenita	/	[17]
37	βadajareu	/	[17]
38	jua nekonə	/	[17]
39	nenerigu	/	[17]
40	tʃauiba tʃutaiba	/	[17]

doi:10.1371/journal.pone.0094814.t001

parentheses in Tables 2–3). Therefore, it is doubtful whether geminates, aspirated consonants, and long vowels really exist in Carabayo. Note that the absence of *s* is confirmed by Juan Berchmans de Felanix's [17] observation that the Carabayo man pronounced Spanish very well, except for *s*, which he pronounced as *ʃ*. Despite the scarcity of the data, Tables 2 and 3 represent what might be a perfectly plausible and also typically Amazonian

phoneme system, suggesting that a comparative analysis can reasonably be carried out with these data.

Can we tell from this material whether the Carabayo language is related in any way to any other known language? One hypothesis would be that they speak a closely related variant of a living language. This appears to be the case in neighboring Peru, where most uncontacted groups are linguistically related to groups already contacted, which allows one to have some degree of

Table 2. Carabayo consonants.

	bilabial		alveolar		palatal	velar		glottal
plosive	b ₆	p ₅	(d ₂)	t ₅		g ₆	k ₁₃	
(plosive geminate)			(dd ₁)					
(plosive aspirated)							(kh ₁)	
fricative	(β ₂)				j ₅	(x ₁)		(h ₂)
affricate					ʃ ₅			
approximant								
nasal	m ₉		n ₂₂		ɲ ₆			
nasal geminate			nn ₃					
Flap			r ₁₁					
(liquid)			(l ₁)					

Subscript numbers represent frequency of occurrence of the phoneme in the corpus (phonemes in parentheses occur only once or twice).
doi:10.1371/journal.pone.0094814.t002

previous knowledge of their language. However, in the case of the Carabayo, this is unlikely because, while the Carabayo family was held at La Pedrera, speakers of the following languages were asked to try to communicate with them, without success [8,20] (language names are given in standard spelling and with genealogical affiliation and ISO 639-3 codes): Andoke (isolate, ano), Muinane (Boran; bmr), Witoto (Witotoan; three varieties: Μηηηca hto; Murui huu; Nüpode hux), Mirañas and Boras (Boran; both boa), Carijona (Cariban; cbd), Yucuna and Matapí (Arawakan; both ycn), Tanimuca (Tucanoan; tnc), Cabiyaí (Arawakan; cbb), Tikuna (Tikuna-Yurí; tca), Ocaina (Witotoan; oca), Nonuya (Witotoan; noj), Puinave (isolate; pui), Tukano (Tucanoan; tuo), Yuhup (Nadahup; yab). These include all living languages spoken in the region where the Yurí were sighted, and we can thus discard this hypothesis (but see our discussion on Tikuna below).

A second hypothesis is that Carabayo corresponds or is closely related to an extinct but documented language of the region. Based on information given by Bartolomé de Igualada & Marcelino de Castellví [22] and Marcelino de Castellví & Espinosa Pérez [23], as well as other historical sources summarized by Franco [16], the languages given in Table 4 are possible candidates for identifying Carabayo since they were reportedly spoken in the same area, some of them up until the early 20th century. All of them were documented in wordlists in the first half of the 19th century, and all of them are presumed to now be extinct (although Mura's dialect – or sister language – Pirahã survives). Three closely related Arawakan languages [24] were reportedly spoken in the region (Uainuma, Jumana, Passé). Among these, Passé and Uainuma were chosen for the comparison since they are the ones geographically closest to where Carabayo were last sighted.

All of the wordlists for the languages listed in Table 4 were published by Martius [25], except for Natterer's [11] Yurí wordlist (see below). The wordlists were collected in 1819 and 1820 by the German botanists Carl Friedrich Philipp von Martius (Coretú, Coëruna, Mura) and Johann Baptist von Spix (Passé). The Yurí and Uainuma wordlists combine words collected by Martius and by Spix as well as words collected by Alfred Russell Wallace around 1850, which Wallace himself also published in an appendix to Wallace [19]. The Austrian naturalist Johann Natterer collected an additional wordlist on Yurí in 1833 [11]. These wordlists cover basic vocabulary and local fauna and flora terms.

Finally, it is also possible that the Carabayo speak a language that has never been documented. In this context it is noteworthy that a number of languages of the area were documented for the first time as late as the early 20th century, among them Ocaina, Nonuya, and Resígaro, showing that some languages remained unnoticed for a long time after the region had begun to be explored. However, during the 1930s, the indigenous groups of the Caquetá-Putumayo area of the Colombian Amazon region were surveyed in great detail by Capuchin missionaries, including the Ocainas, Nonuyas, and Resígaros [22,23]. Based on information from these surveys, Marcelino de Castellví and Espinosa Pérez [23] suggest that Yurí speakers persist in locations very close to where the Carabayo were sighted in 1969, without, however, giving linguistic data as evidence.

Results and Discussion

Our comparison of the Carabayo data with Coëruna, Coretú, Mura, Passé, Uainuma, and Yurí revealed that a number of Carabayo forms match corresponding Yurí elements, but none match forms of the other languages. Among the Carabayo-Yurí correspondences is one that Vidal y Pinell [8] had identified, Carabayo *ao* 'father', as we discuss below. The other two Carabayo-Yurí correspondences given by Vidal y Pinell [8] do not hold up to scrutiny: He suggests that Carabayo *aua*, which according to Juan Berchmans de Felanix [17] might mean 'child', corresponds to Wallace's Yurí *owíye* 'son'. This correspondence seems far-fetched and cannot be confirmed by other Yurí forms such as *o nné* 'son', *o ñn* 'child', *ta ñnna* 'boy' (Natterer), *oná* 'son', *uhé* 'child' (Martius), or *suuné* (Spix). (Incidentally, Wallace's *owíye* 'son' probably means 'daughter', rather than 'son', as the forms for 'daughter' given by the three other sources for Yurí suggest: *zo ábü* (Natterer), *tshöwü* (Martius), *suabüe* (Spix). The first syllable in these three is the first person possessor marker.) Furthermore Vidal y

Table 3. Carabayo vowels.

	e ₂₆	o ₂₂
i ₁₅		
u ₁₇ /(uu ₁)	ə/(œ ₂)	a ₅₂

Subscript numbers represent frequency of occurrence of the phoneme in the corpus (phonemes in parentheses occur only once or twice).
doi:10.1371/journal.pone.0094814.t003

Table 4. Candidate languages for the identification of Carabayo.

Language	Affiliation	Evidence for affiliation
Coëruna	possibly Witotoan	Koch-Grünberg [31], Loukotka [32]
Coretú	Tucanoan	Loukotka [32]
Mura	Mura(-Pirahã)	Campbell & Grondona [33]
Passé	Arawakan	Ramirez [24]
Uainuma	Arawakan	Ramirez [24]
Yurí	Tikuna-Yurí	Carvalho [7], Goulard & Montes Rodriguez [2]

doi:10.1371/journal.pone.0094814.t004

Pinell [8] suggests that the Carabayo form *ñe*, reported to mean ‘no’, corresponds to Wallace’s Yurí *eeñ* ‘bad’. Again, this seems far-fetched and cannot be confirmed by other Yurí forms for ‘no’: *ka* (Natterer), *tivá* (Martius), *ghainà* (Spix).

There are a number of other forms, however, that display intriguing correspondences between Carabayo and Yurí and that were not detected by Vidal y Pinell [8], partially because he did not have access to Martius’, Spix’s, and Natterer’s Yurí data. The relevant Carabayo and Yurí data are presented in Table 5.

Item 1 in Table 5 contains a complex form in which probably both elements correspond. The first element, *tʃau-* is well attested in Yurí as a first person subject and possessor prefix. It appears in various spellings in Yurí data, e.g. *tschau-*, *tschu-* (Martius), *su-* (Spix), and *tcho-* (Wallace). The apparent mismatch between first person subject form and second person reference in item 1 could easily have arisen due to the lack of a common language in the situation in which the form was noted by Juan Berchmans de Felanix [17]: It is common even in professional fieldwork elicitation situations that, for example, in response to a field worker asking for a translation of “I sit”, an informant provides a form meaning “you sit”, referring to the field worker. Alternatively, *tʃau-* in item 1 may be an object pronoun followed by a prefixless imperative verb form in item 1. The second element, *noʃe* ‘warm’ matches well with Wallace’s Yurí *noré* ‘warm’. It matches less well with Natterer’s form for ‘warm’, but within this form *ore* is shared. Item 2 in Table 5 is a less clear case, but it might be argued that a first person singular form is likely to occur in a form translated as ‘like’. The correspondences involving Carabayo *tʃau-*, *tʃu-*, and *tʃe* proposed in items 3–5 in Table 5 are more hypothetical since no information on their meaning in Carabayo is

available. However, they might contain further instances of the word-initial first person singular prefix. In Yurí, variants of *tʃau-*, probably conditioned by the stem to which it is prefixed, are attested, primarily *tʃu-*, e.g. *tschu-báacki* ‘my elbow’ (Martius). The Carabayo words beginning with *tʃau-*, *tʃu-*, and *tʃe* given in items 3–5 might thus well be nouns with a first person singular possessor prefix or verbs with a first person singular subject prefix that Juan Berchmans de Felanix [17] overheard from the conversations among the Carabayo. Note that the occurrence of *tʃau-*, *tʃu-* is also responsible for the high frequency of *tʃ* in both Carabayo and Yurí, which Vidal y Pinell [8] noted.

Item 6 in Table 5, *hono* ‘boy’, constitutes a reasonably certain correspondence in terms of a sequence of a back rounded vowel (*o* or *u*) followed by *n* and possibly another, unidentified vowel, and is attested as such five times in the Yurí data, including attestations from three different sources. Item 7, Carabayo *ao* ‘father’, also matches reasonably well with Yurí data, as already noted by Vidal y Pinell [8], in terms of the initial vowel *a* and final vowel *o*, which alternates with *u* in the Yurí data. A form related to Yurí (*h*)*ato*, *atu* ‘father’ may also be identifiable in Carabayo *jakomanate*, the Carabayo man’s name, when compared with *jakoma*, the Carabayo man’s eldest son’s name, according to Venanci d’Arenys de Mar [21] (items 24 and 25 in Table 1). If one assumes that the first term literally means ‘Jakoma’s father’, then *nate* would mean ‘father’. This form matches attested Yurí forms relatively well, and it is strikingly similar to Tikuna (Yurí’s sister language) *naŋu* ‘father’. The use of teknonyms is not attested for Tikuna or for other indigenous groups in the direct vicinity, but it is attested in other places in Amazonia. In any case, it seems clear that *jakomanate* is a complex form and it is likely that *-nate* means ‘father’, even if the

Table 5. Summary of Carabayo and Yurí data compared.

CARABAYO	YURÍ MATERIAL COMPARED
1 <i>tʃaunoʃe</i> ‘warm me’	<i>tʃau-</i> + <i>noré</i> ‘warm’ (Wallace)/ <i>tso tsórerú</i> ‘warm’ (Natterer)
2 <i>tʃauameni</i> ‘good, well, like’	<i>tʃau-</i> + (<i>su-</i>) <i>mêniko</i> ‘(my) heart’ (Spix)
3 <i>tʃauiba</i>	<i>tʃau-</i> + (<i>tschu-</i>) <i>ibaüh</i> ‘(my) back’ (Martius)
4 <i>tʃutaiba</i>	<i>tʃau-</i> + <i>taobi</i> (Martius) ‘body’/ <i>taiaeboí</i> (Martius) <i>toipuy</i> (Spix) ‘week’/ <i>taiaeboí</i> (Martius), <i>toipuy</i> (Spix), <i>tai rôn i</i> (Natterer) ‘night’
5 <i>arə tʃe o neko</i>	<i>a aré</i> (Natterer), <i>áhre</i> (Martius), <i>aré</i> (Spix), <i>ahri</i> (Wallace) ‘red’ + <i>tschauúnäco</i> ‘I bury’ (Martius)/ <i>tʃau</i> + <i>nihcó</i> ‘live’ (Martius)/ <i>tschu-inicko</i> (Martius), <i>subinigho</i> (Spix) ‘my testicles’
6 <i>hono</i> ‘boy’	<i>o nné</i> ‘son’, <i>o ên</i> ‘child’, <i>ta ünna</i> ‘boy’ (Natterer), <i>oná</i> ‘son’ (Martius), (<i>su</i>) <i>uné</i> ‘(my) son’ (Spix)
7 <i>ao</i> , <i>-nate</i> ‘father’	<i>atú</i> (Natterer), <i>hato</i> (Martius), <i>háto</i> (Wallace), (<i>su</i>) <i>átu</i> ‘(my) father’ (Spix)
8 <i>hako</i> ‘well!’	<i>hokó</i> ‘I am fine, this is good’ (Natterer), <i>ockó</i> (Martius) <i>ukó</i> (Spix) ‘beautiful’

doi:10.1371/journal.pone.0094814.t005

Table 6. Corresponding elements Carabayo-Yuri vs. other candidate languages (NA - no form attested with a comparable meaning).

GLOSS	CARABAYO	YURI	COERUNA	UAINUMA	PASSÉ	CORETÚ	MURA
1 I, my	t/au-, t/u-	t/au-, t/u-	kui-, ku-	no-, nu-	no-, na-	yi-	tschäng
2 warm	noje	nore	(NA)	amoiri	(NA)	(NA)	(NA)
3 boy, son	hono	ona, one	quáda	-iri	aghunghii, tschiker-noma	simagó	oohabäh
4 father	ao	(h)ato, atu	Domú	pai	payü	tsáackö	itshüaeng
5 well	hako	hokó	(NA)	misare	(NA)	(NA)	(NA)

doi:10.1371/journal.pone.0094814.t006

Carabayo do not employ a conventionalized system of teknonyms. Note that Venancy d'Arenys de Mar [21] claims that the Carabayo man called *jakomanate* was not the father of the oldest Carabayo boy, who was called *jakoma*, but maybe his brother, without, however, providing any evidence or further argumentation for this claim. This claim contradicts all other sources, who assume they were father and son, especially Bergès [26], who probably knew the Carabayo best. Even if they were not father and son, they may have used a teknonym since it has been observed elsewhere in Amazonia that teknonyms are applied among relatives or people living together [27].

The Carabayo expression in item 8 in Table 5 is translated in Juan Berchmans de Felanix [17] as 'bites'. However, the context where this word was recorded is described by Juan Berchmans de Felanix [20] as follows: Shortly after the Carabayo family was captured, they were led, bound, through the jungle. When they arrived at a place where the commission had left their dogs behind, the Carabayo family showed fear and repeated various times the word *hako* ("Al llegar al sitio donde estaban los perros, demostraron miedo, repitiendo distintas veces la palabra JACO" [20]). In this context it is possible that *hako* is some kind of interjection, especially because it was repeated various times. If so, it matches well with the Yuri form *hokó* which is given by Natterer as an equivalent of both 'this is good' [German *dies ist gut*] (contrasting with 'this is not good' [German *dies taugt nichts*], the preceding entry in Natterer's list) and 'I am fine' [German *Mir geht es gut*] (apparently as an answer to 'how are you?' [German *Wie geht es dir*], the preceding entry in Natterer's list). Natterer's Yuri *hokó* probably corresponds to Martius' Yuri *ockó* and Spix's Yuri *ukó*, both given as equivalents of 'beautiful' [Latin *pulcher* in the original list]. The fact that it appears in various contexts suggests that Yuri *hokó* is a more widely applicable expression that may be translated as "well" and that can also be used as an interjection rather than a literal translation of the equivalents given by Martius, Spix, and Natterer. Our experience with indigenous people of the area suggests that it is not unlikely that the same expression would be used in the contexts described for Carabayo *hako* as well as in the ones described for Yuri *hokó*, *ockó*, and *ukó*. For instance, the Bora people, the Yuri's neighbors to the west, would use *tehdujuco*, which literally means 'already like this', in all of these contexts.

Additionally, we note that there are a number of further, far more hypothetical correspondences contained in the data summarized in Table 5. Firstly, Carabayo *meni* (item 2) may correspond to Yuri *meniko* 'heart' if one assumes that an expression translated as 'good, well, like' would be expressed as '(pleases) my heart'. Furthermore, in the Carabayo material for which no translation at all is provided, a number of forms can be identified that match Yuri forms, as noted in Table 5. For instance, *arə* in item 5 may correspond to Yuri *are* 'red', which is well attested in various sources for Yuri.

The two Nheengatu elements in Carabayo, *kariba* 'white man' and *tupana* 'God' are also attested in Yuri data: *kalibbā* (Natterer) and *tupana* (Martius). These correspondences do not provide evidence for an identification of Carabayo with Yuri because both items are widespread among languages of the region. However, the exact match between Yuri *tupana* (Martius) and Carabayo *tupana* is noteworthy, given that this form was apparently phonologically nativized differently in Coeruna, as *toibá*, and in Mura, as *tupaua*. For Uainuma, *tupana* is reported, as well. For the other two candidate languages, words for 'God' which are non-related and probably native are documented, i.e. *pokené* for Passé and *nimúpalightär* for Coretú. No forms corresponding to *kariba* are attested in any of the candidate languages, except for Yuri, due to

Table 7. Carabayo-Tikuna correspondences (Abbreviations: SG – singular; PL – plural; # - word boundary; Ø – elision).

	CARABAYO	TIKUNA	SOUND CORRESPONDENCES
1	<i>ʃʃʰaumeni</i> 'good, well, like'	<i>ʃʃʰau na me nii</i> (1SG/3SG/like/be) 'I like it' (lit. 'it is good to me')	Ø-#n
2	<i>gudda</i> 'wait!'	<i>ŋiŋná</i> 'wait!, not yet'	g-ŋ, dd-n, u-ŋ
3	<i>pinə</i> 'shrimp'	<i>pinŋ</i> 'shrimp species (big, lives in creeks)'	ə-ŋ
4	<i>agó</i> 'bring!'	<i>a ŋe</i> 'bring!' (3SG/bring)	g-ŋ, e-o
(= 3+4)	<i>pinə-gó</i> 'bring shrimp!'	<i>pinŋ na ŋe</i> (shrimp/3SG/bring) 'bring shrimp!'	(see 3, 4)
5	<i>gu</i> 'yes'	<i>ŋiŋ</i> 'yes'	g-ŋ, u-ŋ
6	<i>ʃe</i> 'no'	<i>ʃé</i> 'emphatic negation'	e-é
7	<i>-nate</i> 'father'	<i>natŋ</i> 'father'	(exception to Ø-#n)
8	<i>amá</i> 'come!'	<i>ʃiama</i> 'let us follow'	Ø-#ʃi
9	<i>pama</i> 'there, look!'	<i>paamə</i> 'Quick!, Hurry up!'	a-aa
10	<i>ao</i> 'children to call their father'	<i>a o</i> 'children to call their parents or parents to call children (affective)'	
11	<i>ʃla</i> 'out'	<i>na ʃla</i> (3SG/out) 'get out!'	
12	<i>aua</i> 'calling a child'	<i>na ŋā</i> (3SG/go) 'come here!, move!'	Ø-#n
13	<i>ʃla-nauué</i> 'give me, show me'	<i>ʃla na uué</i> (EXHORTATIVE/3SG/lower) 'lower it (e.g., your hand)!'	
14	<i>hako</i> 'bite'	<i>ja ŋo</i> (EXHORTATIVE/eat) 'eat!'; <i>na ŋo</i> (3SG/eat) 'he eats'	(k-ŋ)
15	<i>ʃàma</i> 'enough!'	<i>tama</i> 'negation'; <i>ʃama</i> '1SG'	
16	<i>dimene</i> 'kill'	<i>tŋmaní</i> (1PL/kill/AGENTIVE) 'our killers'; <i>dŋmení</i> 'look!'	

doi:10.1371/journal.pone.0094814.t007

the fact that there were no entries for this concept in the wordlist template that Martius, Spix, and Wallace used.

In summary, we can identify in Carabayo data four forms that match corresponding Yuri forms well: a first person singular prefix, and words for 'warm', 'boy', and 'father', in addition to a more hypothetical correspondence involving an interjection 'well!'. Table 6 contrasts these Carabayo-Yuri correspondences with non-corresponding forms of other candidate languages.

The strongest evidence for a link between Carabayo and Yuri comes from the first person singular prefix (item 1 in Table 6). It is attested in one complex Carabayo form, for which there is a matching translation (*tʃaunoʃe* 'warm me'), and may be contained in further Carabayo forms given by Juan Berchmans de Felanix [17] without translations. For this form, the absence of corresponding forms in other candidate languages is particularly telling since for a first person singular pronoun, the absence of a corresponding form in a candidate language is not likely due to alternative words with similar meanings that happen to be recorded in the extant wordlist (as can easily happen with words for 'warm', 'boy, son', 'father', and 'well!'). Additionally, the first person singular forms of Yuri, Coeruna, Uainuma, Coretú, and Mura given in Table 6 have etymologies in their linguistic families, which excludes the possibility that they are mistakenly given as first person forms in the wordlist collection in the 19th century. And finally, personal pronouns are known to be especially resistant to borrowing [28], i.e. their similarity is truly indicative of a genealogical link, not of language contact.

It should be noted that any of the suggested correspondences given in Table 6 involve a fair amount of speculation due to the scarcity of information on Carabayo as well as on Yuri. For instance, *tʃaunoʃe* recorded in the context "the speaker ordered a child to warm his hands with fire and apply them to his body," could mean many other things, e.g. 'touch me', 'rub me', or 'hug me'. In addition, there is an unexplained correspondence between the bilabial consonant *β* in Carabayo *noʃe* and the alveolar consonant *r* in Yuri *nore* 'warm'. However, the existence of a whole

set of five at least potentially matching forms shared by Carabayo and Yuri, and the lack of any matching forms in other candidate languages does strongly suggest that if Carabayo is related to any of the candidate languages, it is most likely related to Yuri.

Our comparison of Carabayo and Tikuna revealed a high number of very good matches between Tikuna and Carabayo, as summarized in Table 7. The Tikuna correspondences to Carabayo were provided by Abel Antonio Santos Angarita, a native speaker of Tikuna and trained linguist specializing in Tikuna dialectology [29], on inspection of the Carabayo material. These data contain at least 13 close correspondences. Among these, six items (numbers 1–7 in Table 7) constitute very good matches, both semantically and phonologically. Another six items (numbers 8–13 in Table 7) can be considered good matches. Another three items (numbers 14–16) are given here that match less well but are still worth considering (item 14 provides an alternative correspondence for *hako*). What adds credibility to the matches in Table 7 is that they exhibit regular sound correspondences between Carabayo and Tikuna, especially Carabayo *g* (or *k*) – Tikuna *ŋ* and loss of word-initial *n* in Carabayo, both of which are attested in three well-matching pairs (counting also one case of loss of word-initial *ŋ*). The matching elements include a number of items that are cross-linguistically very hard to borrow, especially first and third person pronouns and the verb 'come' [28]. Even for the other items, similarity is unlikely due to contact since there is a strong cultural avoidance of lexical borrowing in the entire region, and it is unlikely that the Carabayo would be an exception.

Thus the correspondences in Table 7 provide strong indications that Carabayo is genealogically related to Tikuna, but they cannot be taken as evidence for a closer relation with Tikuna than with Yuri, as the larger number of correspondences with Tikuna might suggest. In fact, we may expect a lower number of correspondences with the available Yuri data for a number of reasons. Firstly, Yuri is probably poorly represented, both phonologically and semantically, in the 19th-century data by travelers with no training in linguistics and probably no common language with the

Yurí they were interviewing. Secondly, with only a fixed list of Yurí words available, it is naturally much less likely to find matching elements than when a native Tikuna speaker actively searches correspondences to Carabayo items. Indeed, we have initial evidence that Carabayo shares features with Yurí but not with Tikuna, mainly *g* (or *k*) in positions that correspond to Tikuna *ŋ*, e.g. in Yurí *gwo* – Tikuna *ngoó* ‘snake’ or Yurí *kō ja* – Tikuna *ngùga* ‘Tinamus bird’. All this suggests that Yurí, Carabayo and the various dialects of Tikuna are genealogically related, with Carabayo somewhere in the middle between Yurí and Tikuna, but probably closer to Yurí. The ease with which Carabayo data could be interpreted by a native Tikuna speaker additionally suggests that these languages are relatively closely related and may even form – or have formed in the past – a dialect continuum.

Conclusions

This paper presents evidence suggesting that the Carabayo people, who live in voluntary isolation in the Colombian Amazon region, speak a language related to Yurí and also Tikuna, i.e. that they are – direct or indirect – descendants of the Yurís that travelers such as Martius, Spix, Wallace, and Natterer encountered in the 19th century. We were able to provide correspondences to almost all Carabayo items for which reasonably reliable glosses are available. The correspondences we find between Carabayo and Yurí, on the one hand, and Carabayo and Tikuna,

on the other hand, are unlikely to be instances of borrowing from Yurí and Tikuna and thus likely to truly reflect a genealogical link.

With the accelerating loss of indigenous languages, it becomes increasingly difficult to gain any further knowledge of the pre-colonial linguistic landscape of Amazonia. However, our meticulous study of Carabayo data from 1969 contributes to putting one language, Carabayo(-Yurí), (back) on the map, and to placing another language, Tikuna, (back) in a linguistic family, Tikuna-Yurí, of which it had been presumed to be the only surviving member. We hope that this study will also contribute to awareness of the existence of groups that avoid contact and especially of their right to be left in peace.

Acknowledgments

We are grateful to Fra Valentí Serra de Manresa, the director of the *Arxiu Provincial dels Caputxins de Catalunya* in Sarrià, Barcelona, for help with locating the document by Juan Berchmans de Felanix [17], to Harald Hammarström for initially alerting us to the existence of this word list and discussion of the argument presented here, to Willem Adelaar and Hélène B. Brijnen for generously sharing the unpublished Yurí data collected by Natterer, to Abel Antonio Santos for granting us permission to publish his data, to Bernard Comrie for the suggestion that *Jakomanate* may be a teknonym, and to two anonymous reviewers of PLOS ONE.

Author Contributions

Analyzed the data: FS JAE. Wrote the paper: FS.

References

- Jolkesky M (2009) Macro-Daha: reconstrução de um tronco linguístico do noroeste amazônico. Paper presented at ROSAE - I Congresso Internacional de Linguística Histórica, Salvador do Bahia.
- Goulard JP, Montes MER (2013) Los Yurí/Jurí-Tikuna, en el complejo socio-lingüístico del noroeste amazónico. LIAMES - Línguas Indígenas Americanas 13: 7–65.
- Lewis MP, Simons GF, Fennig CD, editors (2013) Ethnologue: Languages of the World, Seventeenth edition. Dallas, Texas: SIL International. Available: <http://www.ethnologue.com>. Accessed 2013 June 13.
- Rivet P (1912) Affinités du Tikuna. Journal de la Société des Américanistes 9: 83–110.
- Nimuendajú C (1952) The Tukuna. Berkeley: University of California Press.
- Nimuendajú C (1977) Os índios tukuna. Boletim do Museu do Índio, Antropologia 7: 20–68.
- De Carvalho FO (2009) On the genetic kinship of the languages Tikúna and Yurí. Revista Brasileira de Linguística Antropológica 1: 247–268.
- Vidal y Pinell R (1970) Identificación de la tribu de los Yuríes en el Amazonas de Colombia. Amazonia Colombiana Americanista 7: 95–109.
- Trupp F (1974) Una Tribu Desconocida en la Amazonia Colombiana. Bulletin of the International Committee on Urgent Anthropological and Ethnological Research 16: 109–110.
- Landaburu J (2000) Clasificación de las lenguas indígenas de Colombia. In: Gonzáles de Pérez MS, Rodríguez de Montes ML, editors. Lenguas indígenas de Colombia. Una visión descriptiva. Bogotá: Instituto Caro y Cuervo. pp. 25–48.
- Natterer J (1833) Sprachproben der Nation Jurí – aus dem Rio Jupurá – (Schupurá). Manuscript held at the University Library of Basle.
- Anonymous (probably Antonio Font [Fray Juan Berchmans de Felanix]) (1969) Editorial: Donde está la falla??? Amanecer Amazónico XVI: 1.
- Hammarström H (2010) The status of the least documented language families in the world. Language Documentation & Conservation 4: 177–212.
- Gottron T, Lipka N (2010) A Comparison of Language Identification Approaches on Short, Query-Style Texts. In: Gurrin C, He Y, Kazai G, Kruschwitz U, Little S, et al., editors. Advances in Information Retrieval. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer. pp. 611–614. Available: http://link.springer.com/chapter/10.1007/978-3-642-12275-0_59. Accessed 2 January 2014.
- Durie M, Ross M, editors (1996) The Comparative Method Reviewed. Regularity and Irregularity in Language Change. Oxford: Oxford University Press.
- Franco R (2012) Cariba malo: episodios de resistencia de un pueblo indígena aislado del Amazonas. Leticia: Universidad Nacional de Colombia. Available: www.bdigital.unal.edu.co/6140/3/9789587611618.pdf.
- Juan Berchmans de Felanix (1969) Nuestros indígenas de La Pedrera. Amanecer Amazónico XVI: 5, 16–17.
- Lexicon Cappucinum (1951) Rome: Collegio Internaz. S. Lorenzo da Brindisi.
- Wallace AR (1853) A narrative of travels on the Amazon and Rio Negro, with an account of the native tribes, and observations on the climate, geology, and natural history of the Amazon Valley. London: Reeve.
- Juan Berchmans de Felanix (1969) Relato auténtico de la expedición a la tribu desconocida en busca de Julián Gil y Alberto Miraña desaparecidos a fines de enero del corriente año de 1969 -en plena selva de la Amazonia Colombiana. Separata de Amanecer Amazónico 744: 1–9.
- Venanci d'Arenys de Mar [Xavier Nogueras i Prats] (1969) Descobriments d'una tribu totalment primitiva a la Missió dels Caputxins catalans de Leticia (Amazonas - Colombia, S.A.) I-IV. Igualada: Gráficas Argent [Relaciones fechadas en Nazaret (Amazonas) el 13 de mayo, 4 de junio, 15 de junio y en Barcelona Sarrià el 17 de septiembre de 1969 y publicada por separata y en “El Apostolado Franciscano”. Barcelona, Año 60, 1969, set-oct., 137–142, 158–164, 182–188; año 1970, feb., 40–46.].
- Bartolomé de Igualada, Marcelino de Castellví (1940) Clasificación y estadística de las lenguas habladas en el Putumayo, Caquetá y Amazonas. Amazonia Colombiana Americanista. Órgano semestral de materiales y estudios para la Americanística del CILEAC de Sibundoy 1: 91–101.
- Marcelino de Castellví, Espinosa Pérez L (1958) Propedéutica etnolingüística y diccionario clasificador de las lenguas indoeuropeas. Madrid: Consejo Superior de Investigaciones Científicas.
- Ramírez H (2001) Línguas Arawak da Amazônia Setentrional: Comparação e Descrição. Manaus: Editora de Universidade do Amazonas.
- Martius CFP von (1867) Beiträge zur Ethnographie und Sprachkunde Amerika's, zumal Brasiliens. Vol. II: Zur Sprachkunde. Wörtersammlung Brasilianischer Sprachen. Glossaria linguarum Brasiliensium. Glossarios de diversas línguas e dialectos, que fallao os Índios no império de Brazil. Leipzig: Friedrich Fleischer.
- Bergès Y-G (1970) La lune est en Amazonie. Paris: A. Michel.
- Vilaça A (2002) Making Kin out of Others in Amazonia. The Journal of the Royal Anthropological Institute 8: 347–365.
- Tadmor U (2009) Loanwords in the world's languages: Findings and results. In: Haspelmath M, Tadmor U, editors. Loanwords in the World's Languages. A Comparative Handbook. Berlin: Mouton de Gruyter. pp. 55–75.
- Santos AA (2005) Hacia una dialectología Tikuna del Trapecio Amazónico colombiano [M.A. Thesis]. Leticia: Universidad Nacional de Colombia, Sede Amazonia.
- Castro Caycedo G (1978) Perdido en el Amazonas. Bogotá: C. Valencia Editores.
- Koch-Grünberg T (1906) Les Indiens Ouitotos, étude linguistique. Journal de la Société des Américanistes 3: 157–189.
- Loukotka C (1968) Classification of South American Indian languages. Los Angeles: Latin American Center, University of California.
- Campbell L, Grondona V, editors (2012) The Indigenous Languages of South America: A Comprehensive Guide. Berlin, Boston: de Gruyter Mouton.